OHIO RIVER BASIN PRECIPITATION FREQUENCY STUDY

Update of *Technical Paper No. 40*

Fifth Progress Report
for the Period
March 1, 2000 through September 29, 2000

Lesley Tarleton Julian
Bingzhang Lin, Eloisa Raynault, David Riley, Deborah Todd and Michael Yekta

Office of Hydrology
U.S. National Weather Service
Silver Spring, Maryland

September 2000

DISCLAIMER

The data and information presented in this report should be considered as preliminary and are provided only to demonstrate current progress on the various technical tasks associated with this project. Values presented herein are NOT intended for any other use beyond the scope of this progress report. Anyone using any data or information presented in this report for any purpose other than for what it was intended does so at their own risk.

TABLE OF CONTENTS

OHIO RIVER PRECIPITATION FREQUENCY STUDY
Update of *Technical Paper No. 40*

Fifth Progress Report
for the Period
March 1, 2000 through September 29, 2000

I. STUDY OVERVIEW

A. Purpose and Scope

The Hydrometeorological Design Studies Center (HDSC), Office of Hydrology, U.S. National Weather Service is performing a precipitation frequency study to update Technical Paper No. 40, *Rainfall Frequency Atlas of the United States for Durations from 30 Minutes to 24 Hours and Return Periods from 1 to 100 Years* (Hershfield, 1961), for the Ohio River basin. The study involves the completion of specific tasks including collecting and performing quality control of data, compiling and formatting datasets for analyses, selecting applicable frequency distributions and fitting techniques, analyzing data, mapping and preparing reports and other documentation.

B. Study Area

The study area covers 13 states completely and parts of nine additional bordering states. The Susquehanna River and Delaware River basins are also included in the study area.

The study area is divided into 16 near-homogeneous climatic (i.e., an extreme precipitation climate) regions. Factors considered in defining the regions include 1) the season(s) of highest precipitation, 2) the type of precipitation (e.g., general storm, convective, tropical storms or hurricanes, or a combination), 3) the climate, 4) the topography and 5) the homogeneity of these factors in a single area.

The study area is displayed in Figure 1. The core and border states and regional boundaries are shown on the figure.

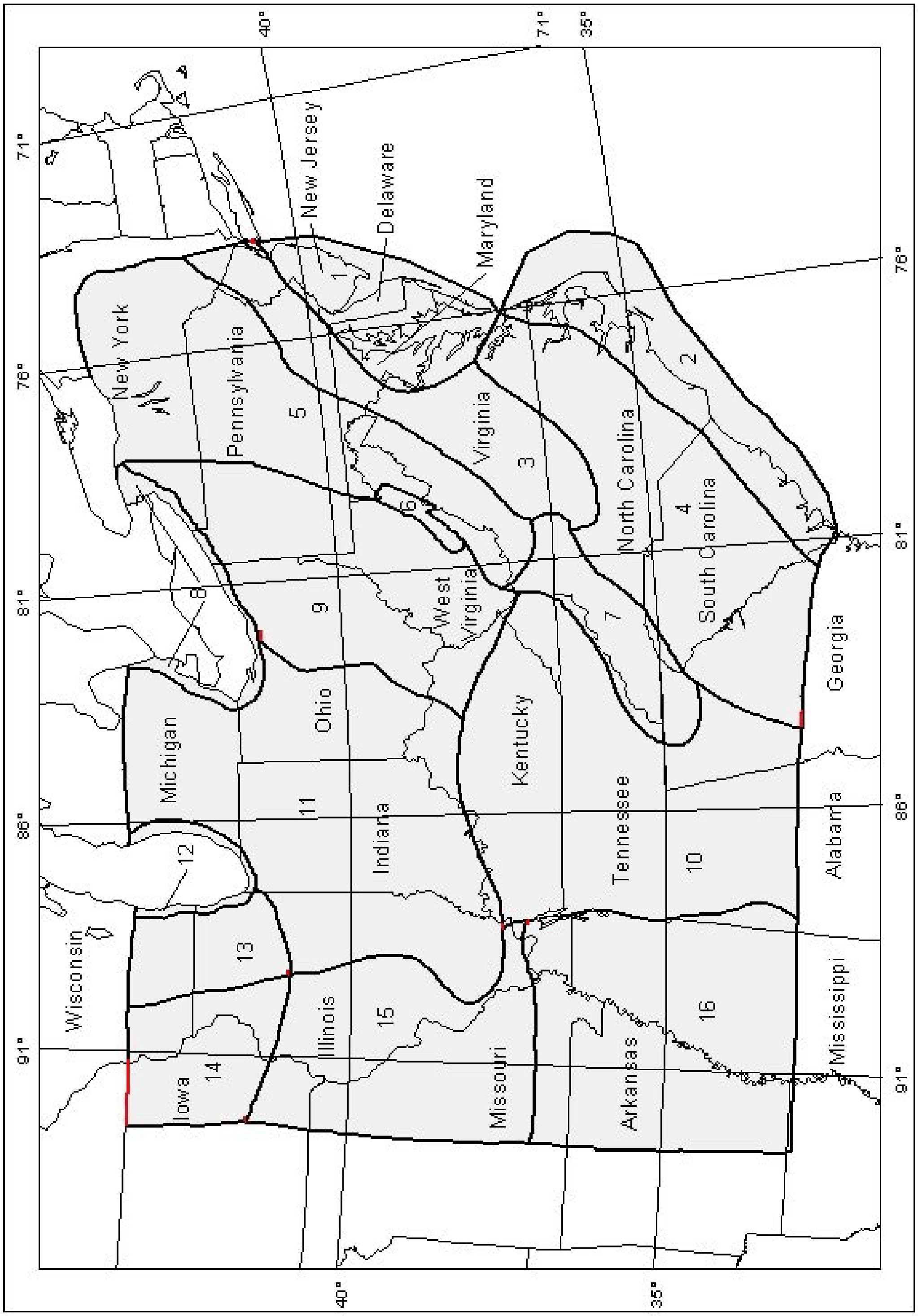


Figure 1. Ohio River study area and region boundaries.

II. TASK STATUS

The following sections discuss the status of each task and provide a short technical description of decisions made and tasks accomplished.

A. Data Collection and Quality Control

The daily, hourly and n-minute datasets are complete. As reported in the Fourth Progress Report, the 15-minute dataset is unusable since it is mainly hourly data in a 15-minute format. The compilation for the daily, hourly and n-minute datasets includes quality control, merging stations where appropriate and formatting for analysis. The criteria for merging stations have been revised. For merging, stations must: 1) be within 100 feet in elevation, 2) be within 5 miles in distance, 3) contain a gap between records of 60 months (5 years) or less, and 4) contain an overlap in records of 60 months (5 years) or less.

If stations have 20 years or more of data, they will be used in the L-moment analyses. If stations have less than 20 years of data, they are available for storm analysis or other investigation. The following sections discuss the status of ongoing or completed activities involving daily, hourly and n-minute datasets.

1. Daily Data

The daily database from the National Climatic Data Center (NCDC), the U.S. Army Corps of Engineers (COE) and the U.S. Geological Survey (USGS) were updated through November 1998. The updating includes quality control, elimination of duplicate records and merging stations for longer station records where appropriate.

Table 1 shows the number of daily stations in each state with merged data. The data are from one or more of the following sources: NCDC, COE and USGS. 182 out of 5512 daily stations have merged records. Because the criteria for merging stations have been modified, any previously merged stations are considered invalid. The 278 out of 5422 daily stations with merged records shown in the Fourth Progress Report may be disregarded.

Table 1. Daily stations with merged data from NCDC, COE and/or USGS. Core states in **bold**.

STATE	No.
Alabama	8
Arkansas	3
Delaware	**0**
Georgia	1
Iowa	2
Illinois	**17**
Indiana	**11**
Kentucky	**22**
Maryland	**8**
Michigan	7
Missouri	3
Mississippi	3
North Carolina	**18**
New Jersey	**8**
New York	10
Ohio	**15**
Pennsylvania	**14**
South Carolina	**9**
Tennessee	**10**
Virginia	**6**
Wisconsin	0
West Virginia	**7**
TOTAL	182

Table 2a is a list of daily stations for L-moment analysis ($\geq$ 20 years of record) from all sources: NCDC, COE and USGS. For the Ohio River Basin area the overall average record length is 52 years. Table 2a also shows the number of stations and average record length for each state. Some station records contain data from more than one source.

Table 2a. Daily dataset including average years using source data. Core states in **bold**.

STATE	No. STATIONS $\geq$ 20 YEARS RECORD	AVERAGE YEARS
Alabama	103	46
Arkansas	86	55
Delaware	**10**	**58**
Georgia	111	52
Iowa	83	60
Illinois	**229**	**62**
Indiana	**181**	**58**
Kentucky	**216**	**50**
Maryland	**93**	**49**
Michigan	74	66
Missouri	145	53
Mississippi	87	55
North Carolina	**255**	**50**
New Jersey	**90**	**55**
New York	212	46
Ohio	**249**	**59**
Pennsylvania	**310**	**48**
South Carolina	**108**	**59**
Tennessee	**253**	**42**
Virginia	**205**	**46**
Wisconsin	57	59
West Virginia	**151**	**47**
TOTALS	3308	52

Table 2b is a list of all daily stations from NCDC, COE and USGS, even those with shorter records. Stations with less than 20 years of record are used for storm analysis and/or more detailed studies.

Table 2b. Daily dataset using source data. Core states in **bold**.

STATE	TOTAL No. STATIONS	No. STATIONS ≤ 20 YEARS RECORD
Alabama	158	55
Arkansas	130	44
Delaware	**19**	**9**
Georgia	147	36
Iowa	148	65
Illinois	**387**	**158**
Indiana	**325**	**144**
Kentucky	**385**	**169**
Maryland	**171**	**78**
Michigan	138	64
Missouri	197	52
Mississippi	111	24
North Carolina	**370**	**115**
New Jersey	**122**	**32**
New York	378	6
Ohio	**394**	**145**
Pennsylvania	**546**	**236**
South Carolina	**149**	**41**
Tennessee	**515**	**262**
Virginia	**330**	**125**
Wisconsin	87	30
West Virginia	**305**	**154**
TOTALS	5512	2204

2. Hourly Data

Hourly data from the NCDC, the COE and the USGS were compiled through November 1998. This includes quality control, elimination of duplicate records and merging stations for longer station records. For merging, stations must: 1) be within 100 feet in elevation, 2) be within 5 miles in distance, 3) contain a gap between records of 60 months (5 years) or less, and 4) contain an overlap in records of 60 months (5 years) or less.

The hourly data obtained from the COE have been reformatted and appended to the NCDC hourly stations for periods with a gap in the dataset. Some COE hourly stations are co-located with daily stations. If a co-located hourly station had a longer record than the daily, the COE hourly values were recomputed as 24-hour (daily) totals and the additional years appended to daily station records.

Table 3 shows the number of hourly stations merged in each state. 110 out of 1940 hourly stations have merged records. Because the criteria for merging stations have been modified, any previously merged stations are considered invalid. The 246 out of 1801 hourly stations with merged records shown in the Fourth Progress Report should be disregarded.

Table 3. NCDC hourly stations with merged data from NCDC, COE and/or USGS. Core states in **bold**.

STATE ID	No.
Alabama	2
Arkansas	3
Delaware	**0**
Georgia	2
Iowa	3
Illinois	**16**
Indiana	**13**
Kentucky	**11**
Maryland	**0**
Michigan	2
Missouri	0
Mississippi	2
North Carolina	**8**
New Jersey	**4**
New York	2
Ohio	**19**
Pennsylvania	**10**
South Carolina	**0**
Tennessee	**2**
Virginia	**3**
Wisconsin	3
West Virginia	**5**
TOTAL	110

Table 4a is a list of hourly stations with ≥ 20 years of record from NCDC, COE and USGS. For the study area the average record length for hourly stations for L-moment analysis (≥ 20 years) is 38 years.

Table 4a. Hourly dataset including average years. Core states in **bold**.

STATE	No. STATIONS ≥ 20 YEARS RECORD	AVERAGE YEARS
Alabama	19	38
Arkansas	25	40
Delaware	**3**	**36**
Georgia	34	40
Iowa	26	40
Illinois	**80**	**40**
Indiana	**76**	**39**
Kentucky	**61**	**38**
Maryland	**16**	**34**
Michigan	24	40
Missouri	49	36
Mississippi	26	44
North Carolina	**51**	**40**
New Jersey	**22**	**36**
New York	47	39
Ohio	**106**	**38**
Pennsylvania	**138**	**36**
South Carolina	**25**	**43**
Tennessee	**48**	**39**
Virginia	**50**	**35**
Wisconsin	16	40
West Virginia	**42**	**40**
TOTALS	984	38

Table 4b is a list of all hourly stations from NCDC, COE and USGS, even those with shorter records. Stations with less than 20 years of record are used for storm analysis and/or more detailed studies.

Table 4b. Hourly dataset using source data. Core states in **bold**.

STATE	TOTAL No. STATIONS	No. STATIONS ≤ 20 YEARS RECORD
Alabama	33	14
Arkansas	35	10
Delaware	**4**	**1**
Georgia	69	34
Iowa	46	20
Illinois	**147**	**67**
Indiana	**136**	**60**
Kentucky	**130**	**69**
Maryland	**42**	**26**
Michigan	45	21
Missouri	81	32
Mississippi	45	19
North Carolina	**111**	**60**
New Jersey	**49**	**27**
New York	101	54
Ohio	**187**	**81**
Pennsylvania	**282**	**144**
South Carolina	**44**	**19**
Tennessee	**143**	**95**
Virginia	**103**	**53**
Wisconsin	22	6
West Virginia	**85**	**43**
TOTALS	1940	956

3. Conversion Factors

Conversion factors were determined based on comparisons of concurrent daily-hourly-minute precipitation data at co-located first-order stations in the Ohio River Basin.

For the daily conversion factors, 86 first order stations had monthly maxima data available to compute conversion factors. The 86 stations were used to compute the conversion factors for 1-day to 24-hour data and for 2-day to 48-hour data.

For the hourly conversion factors, 69 first order stations were co-located with hourly and n-minute stations and used to compute the conversion factor for 1-hour data to 60-minute data. The conversion factor for 2-hour data to 120-minute data was also computed using 68 of the co-located stations.

All relevant data in the project were and will be adjusted by both daily and hourly factors. The conversion factors determined are listed below.

Table 5. List of conversion factors.

Daily Conversion Factors	1-day to 24-hour	***1.13***
	2-day to 48-hour	***1.04***
	No. stations	**86**
	Minimum years of record	**15**
Hourly Conversion Factors	1-hour to 60-minute	***1.16***
	2-hour to 120-minute	***1.05***
	No. stations	**69 for 1-hour** **68 for 2-hour**
	Minimum years of record	**15**

The 1-day to 24-hour conversion factor of 1.13, computed explicitly for the Ohio River Basin Area, is the same as developed in TP40 (1961). The 1-hour to 60-minute conversion factor of 1.16, computed explicitly for the Ohio River Basin Area, is different from the 1.13 value in TP40 (1961). No factors for 2-day to 48-hour and 2-hour to 120-minute were given in TP40.

B. Frequency Distribution Fitting Analyses

As reported in the Fourth Progress Report, this task evaluates and selects the frequency distribution which provides the best fit for the data. For this project the best fit to the partial duration series data, both daily and hourly, is the Generalized Normal distribution (GNO) for precipitation frequency estimates. The best fit selection was based on L-moment analyses by Bingzhang Lin (see Appendix).

C. Precipitation Frequency Value Calculations

The purpose of this task is to obtain a consistent set of precipitation-frequency estimates and relations. The procedure includes defining near-homogeneous regions, and initially 16 regions had been defined (see Figure 1). The Ohio River basin has fairly similar conditions over large areas, and further examination (i.e., discordancy testing, etc.) Indicated no need to change the original regions. The following discusses some of the aspects of L-moment procedures, and their application to regionalization.

L-moment analysis.

L-moment statistics are used for quality-control and return frequency estimates (Hosking and Wallis 1997). L-moment definitions and tests for Heterogeneity (H), Goodness-of-fit and Discordancy (D) are described here.

<u>Heterogeneity.</u>

The heterogeneity test consists of three parts: H-1 based on L-coefficient-of-variation (L-CV), H-2 based on L-CV and L-skewness (L-SK) and H-3 based on L-SK and L-kurtosis (L-KT). Earlier studies (Hosking and Wallis (1991) and conversations with Wallis (1993)) indicate that a threshold of 2 is reasonable, especially for precipitation data. Therefore, for each H-test, a value greater than 2 (H>2) indicates heterogeneity, rather than homogeneity (H<2). As precipitation data are highly variable in any case, the heterogeneity results were considered giving less weight to the L-CV criterion (see Appendix).

<u>Goodness-of-fit.</u>

The approach for goodness-of-fit has been refined. Three methods were used in the selection of the best distribution in the precipitation frequency study (see Appendix).

1. Xtest: This test measures the “distance” of the regional average L-kurtosis, referring to the regional average L-skewness, from various theoretical probability distributions. Monte Carlo simulation was used to obtain a standard deviation of the

"distance" for comparison. The threshold of pass-or-fail for goodness-of-fit tests is 1.64 (absolute value, 90% confidence). The smaller the distance, the better the fit is in terms of goodness-of-fit.

2. Graphical test (4 criteria): This test measures the distances of the L-kurtosis for all real data points in the region from various theoretical distributions. The best fit was based on four criteria: 1) the distance of the average L-kurtosis; 2) the absolute mean distance; 3) the average of absolute distances; 4) the mean-square-root of the distances. No simulation took place.

3. Real-data-check: This test compares the empirical frequencies of the real data with the probabilistic quantiles for various theoretical distributions fitted to the data over all stations (Lin and Vogel 1993).

Based on the three methods, the Generalized Normal (GNO) distribution was selected as the best fit to the Partial Duration (PD) precipitation data for both daily and hourly in the Ohio River Basin.

Discordancy.

The discordancy measure is used for data checking and quality control. In evaluating regions, it is also used to determine if a site had been assigned to the appropriate region. The measure is based on L-moments, specifically L-CV, L-SK and L-KT, which represent a point in 3-dimensional space, for each site. Discordancy is then a function of the distance from the cluster center of points for the sites in the given region. The cluster center is the unweighted mean of the three moments for the sites within the region being tested (see Appendix).

Sites with a discordancy value of 3 or greater are considered discordant (D), while sites with a value of 5 or greater are highly discordant. The latter sites were examined to determine if data problems existed or if they belong in another region. The threshold values of 3 and 5 are not based on a rigorous test, but rather considered reasonable levels to be expected within a homogeneous region.

In the 16 regions, 147 stations were discordant according to the L-moments 1-day partial duration analysis. Of the 147 discordant stations, 111 had discordancy values greater than or equal to 3 and less than 5. The remaining 36 stations had discordancy values greater than or equal to 5. Discordant stations fell into two categories: stations with missing years of data and stations with questionable data values. Daily stations with missing data years were examined using the annual maximum databases.

Of the 36 stations with discordancy values of 5 or greater, eight were entirely deleted from the datasets. Each of the eight stations had its length of record, proximity to other stations and questionable data values examined. Table 6 lists the station identification, latitude, longitude, state, region and discordancy related to each of the eight deleted stations.

Table 6. Deleted discordant (D≥5) stations.

Station ID	**Lat.**	**Lon.**	**State**	**Region**	**Discordancy**
36-0821	40.15	79.03	Pennsylvania	5	13.86
33-3144	41.67	82.82	Ohio	8	5.01
44-4185	37.42	82.02	Virginia	9	11.17
40-4706	36.25	83.60	Tennessee	10	5.29
20-1133	43.20	83.05	Michigan	11	5.06
20-2851	43.02	83.70	Michigan	11	9.95
47-5677	43.00	89.77	Wisconsin	13	5.29
47-0890	42.80	89.87	Wisconsin	14	5.66

Table 7 shows the L-moment test results for Discordancy, Heterogeneity and Goodness-of-fit for daily (1-day) data.

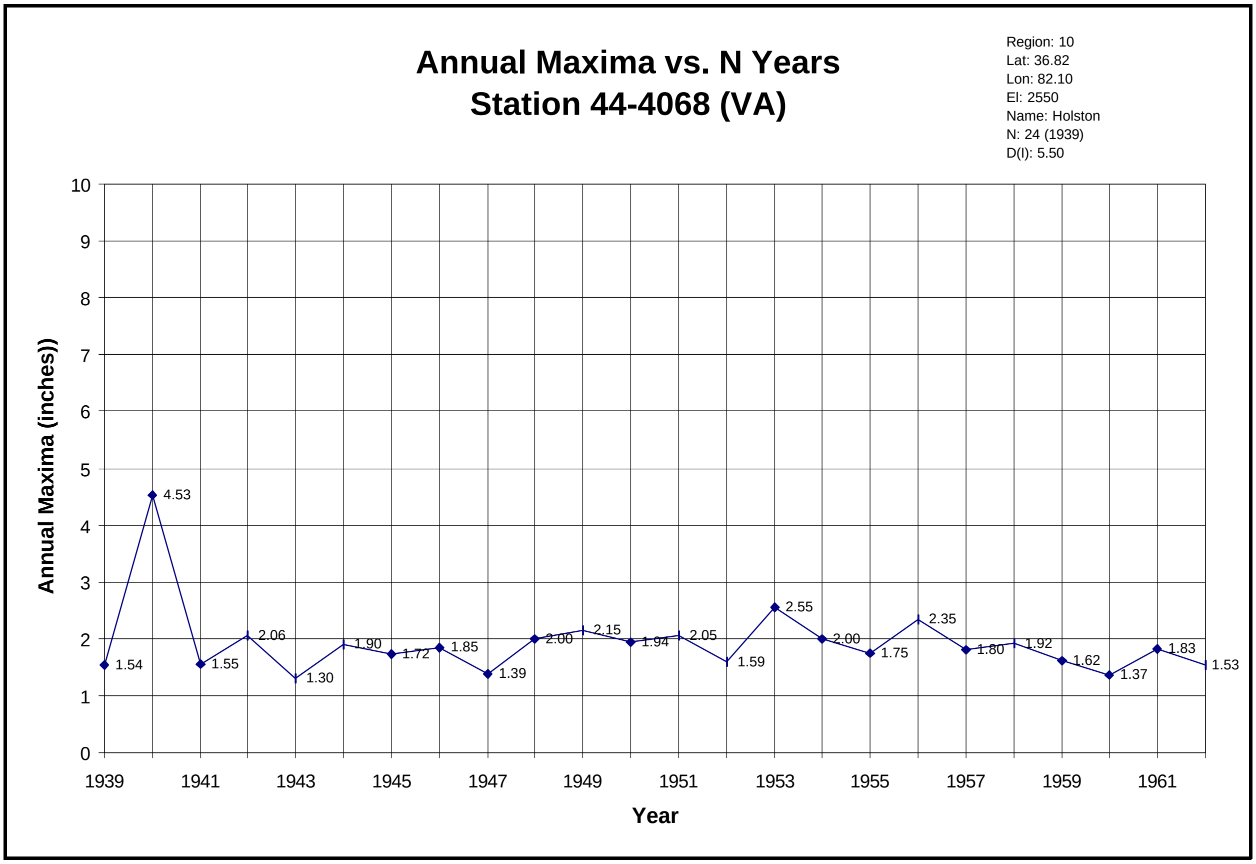
Table 7. Partial Duration Precipitation Series for a 1-day duration.

		Discordancy			Heterogeneity			Goodness-of-fit		
Region	Sites	D_{Total}	$3 \leq D < 5$	$D \geq 5$	H-1	H-2	H-3	GLO	GEV	GNO
1	121	7	3	4	-1.63	-2.31	-3.31	5.73	3.30	**-0.48**
2	46	3	2	1	-2.18	-2.64	-2.71	3.03	**1.47**	**-0.86**
3	203	7	4	3	-1.41	-3.54	-3.29	5.58	2.59	-2.23
4	220	11	5	6	-1.86	-3.20	-3.20	7.98	4.02	**-1.31**
5	412	20	9	11	-0.88	-4.71	-5.90	6.86	2.47	-3.96
6	10	0	0	0	-0.45	-1.78	-2.06	**-1.03**	**-1.40**	-2.18
7	179	8	5	3	-1.22	-3.91	-4.17	4.16	**0.73**	-3.03
8	31	1	0	1	-1.20	-2.39	-2.79	**0.67**	**-0.28**	-2.23
9	324	17	9	8	-1.94	-4.13	-3.96	6.48	2.28	-3.84
10	465	26	19	7	-3.39	-6.13	-6.68	6.07	**0.56**	-5.91
11	351	21	11	10	-1.28	-3.11	-3.79	11.67	6.39	**-0.78**
12	26	1	1	0	-0.82	-0.77	-0.90	**1.55**	**0.46**	**-1.42**
13	68	3	0	3	-0.25	-1.31	-1.24	2.65	0.65	-2.09
14	89	4	1	3	-2.54	-2.95	-2.96	6.35	3.82	**0.35**
15	203	10	3	7	-1.19	-3.76	-4.20	9.20	5.14	**0.05**
16	264	8	3	5	-2.75	-5.43	-5.37	6.16	1.80	-3.27
Totals	3012	147	75	72						

Note:
D_{Total} = Total number of discordancies.
$3 \leq D < 5$ = Number of discordancies greater than or equal to 3 and less than 5.
$D \geq 5$ = Number of discordancies greater than or equal to 5.
H-1 = Based on L-coefficient-of-variation only.
H-2 = Based on L-coefficient-of-variation and L-skewness.
H-3 = Based on L-skewness and L-kurtosis.
GLO = Generalized Logistic distribution.
GEV = Generalized Extreme Value distribution
GNO = Generalized Normal distribution.

By closely examining the time series of the annual maxima for each station that had a discordancy value greater than 5, stations with abnormal data could be identified. The following example in Graph 1 shows a high annual maximum, 4.53 inches, in 1940 at the Holston station (44-4068) in Virginia. Data from surrounding stations at the same time also had high values thus verifying the likelihood of such a high value. Therefore, the station was retained and the data were unchanged though it had a discordancy of 5.50.

Graph 1. Annual Maxima vs. N Years for station 4068 in Virginia.



For 6-hour PD analysis, 33 stations were discordant according to the L-moments results. Of the discordant stations, 24 had discordancy values greater than or equal to 3 and less than 5. Nine remaining stations had discordancy values greater than or equal to 5. After examination, all nine stations were accepted as valid. Table 8 shows the L-moment test results for 6-hour data.

Table 8. Partial Duration Precipitation Series for a 6-hour duration.

		Discordancy			**Heterogeneity**			**Goodness-of-fit**		
Region	Sites	D_{Total}	$3 \leq D < 5$	$D \geq 5$	H-1	H-2	H-3	GLO	GEV	GNO
1	25	2	2	0	-0.07	-1.57	-1.34	1.82	**0.43**	**-0.88**
2	11	0	0	0	0.65	-0.52	-0.32	**0.80**	**0.05**	**-0.88**
3	54	2	2	0	-1.53	-2.90	-3.23	**1.17**	**-0.35**	-2.34
4	54	1	0	1	-1.76	-3.22	-3.47	**1.41**	**-0.20**	-2.33
5	148	6	4	2	-1.72	-4.87	-5.21	**-0.58**	-2.85	-5.86
6	3	0	0	0	-0.40	0.47	0.45	**-0.28**	**-0.60**	**-0.99**
7	35	1	1	0	-0.84	-1.87	-1.92	4.16	**0.73**	-3.03
8	11	0	0	0	-1.16	-1.16	-0.83	**1.28**	**0.59**	**-0.40**
9	137	5	3	2	-1.69	-3.55	-3.77	**1.24**	**-1.23**	-4.34
10	112	4	4	0	-2.11	-3.69	-4.24	**0.72**	**-1.35**	-4.15
11	173	6	4	2	-2.26	-3.82	-4.18	3.49	**0.66**	-2.98
12	18	0	0	0	-1.53	-1.38	-1.19	**0.67**	**-0.13**	**-1.27**
13	17	1	1	0	-1.17	-2.34	-2.46	1.72	**0.51**	**-0.73**
14	36	1	0	1	-1.25	-2.07	-2.29	**0.97**	**-0.25**	-1.98
15	70	1	1	0	-1.09	-1.70	-2.10	2.38	**0.34**	-2.02
16	80	3	2	1	-1.27	-3.07	-2.88	**0.84**	**-1.20**	-3.52
Totals	984	33	24	9						

Note: Refer to Table 7 for abbreviations.

Trend Analysis.

Analysis of time series for trends or data shifts is part of each study. Tests are run for randomness, linear trends and shifts in mean and variance. The Update of TP40 for the Ohio River basin shows that about 85 percent of nearly 2000 daily stations (60 years) have no linear trend in mean or variance. However, for stations with a significant trend, the trend is steadily upward in nearly all cases. Where the trend is downward, it is steadily downward. The shift test for pre-1958 data compared with 1959-1998 has a much stronger signal for extreme precipitation changes. Although fewer stations were available for the shift test, 15.2 percent showed a significant shift upward of the mean, and 54.9 percent showed a significant increase in variance.

Regionalization.

The precipitation frequency estimates are developed from L-moment regional analyses. The L-moment analysis computes regional growth factors (RGFs) for the various return frequencies for each region. As the differences in standard deviations for the RGFs over the study area are relatively small, with standard deviations of 0.11 or less, even at 100-year return frequencies, it was decided that the original 16 regions will not be changed. Table 9 shows the RGFs for the 24-hour (daily and hourly) data.

Table 9. Regional Growth Factors.

Regional Growth Factors									
Region	**2-yr**	5-yr	10-yr	25-yr	50-yr	**100-yr**	200-yr	500-yr	1000-yr
1	**0.90**	1.16	1.37	1.71	2.00	**2.33**	2.70	3.26	3.74
2	**0.90**	1.15	1.36	1.69	1.97	**2.29**	2.64	3.18	3.64
3	**0.91**	1.14	1.35	1.67	1.94	**2.26**	2.61	3.15	3.61
4	**0.92**	1.14	1.33	1.61	1.85	**2.13**	2.43	2.88	3.27
5	**0.92**	1.14	1.32	1.61	1.85	**2.13**	2.45	2.91	3.32
6	**0.91**	1.13	1.33	1.65	1.93	**2.27**	2.65	3.24	3.76
7	**0.93**	1.14	1.31	1.55	1.76	**2.00**	2.25	2.63	2.94
8	**0.91**	1.12	1.32	1.63	1.91	**2.23**	2.60	3.17	3.68
9	**0.92**	1.13	1.30	1.57	1.80	**2.05**	2.34	2.78	3.15
10	**0.93**	1.13	1.30	1.54	1.75	**1.98**	2.24	2.62	2.94
11	**0.92**	1.13	1.32	1.59	1.82	**2.08**	2.38	2.81	3.19
12	**0.91**	1.14	1.33	1.64	1.91	**2.22**	2.57	3.10	3.57
13	**0.91**	1.14	1.34	1.63	1.88	**2.16**	2.47	2.94	3.34
14	**0.91**	1.15	1.35	1.64	1.90	**2.19**	2.51	2.99	3.40
15	**0.92**	1.14	1.32	1.60	1.83	**2.09**	2.38	2.81	3.18
16	**0.93**	1.13	1.30	1.55	1.77	**2.00**	2.27	2.65	2.98
Mean	**0.91**	1.14	1.33	1.62	1.87	**2.15**	2.47	2.95	3.36
S.D.	**0.01**	0.01	0.02	0.05	0.08	**0.11**	0.15	0.22	0.28
Max.	**0.93**	1.16	1.37	1.71	2.00	**2.33**	2.70	3.26	3.76
Min.	**0.90**	1.12	1.30	1.54	1.75	**1.98**	2.24	2.62	2.94
Range	**0.03**	0.03	0.08	0.16	0.25	**0.34**	0.46	0.64	0.82

III. INTERNET-BASED GRAPHICAL USER INTERFACE (GUI)

Work on the Internet-based GUI, which provides point and areal (up to 400 square miles) precipitation frequency data, is nearing completion. Designed after the on-line Alabama rainfall atlas (http://bama.ua.edu/~rain), the HDSC GUI was developed at the University of Alabama. An internal review process for the Java-coded GUI is in progress. Though initiated for the Semiarid Southwest study, the GUI has been designed to accommodate future studies, such as the Ohio River and the Puerto Rico Precipitation Frequency Study. Figure 2 shows the HDSC GUI home page, as it exists in the draft stage. This page is password protected.

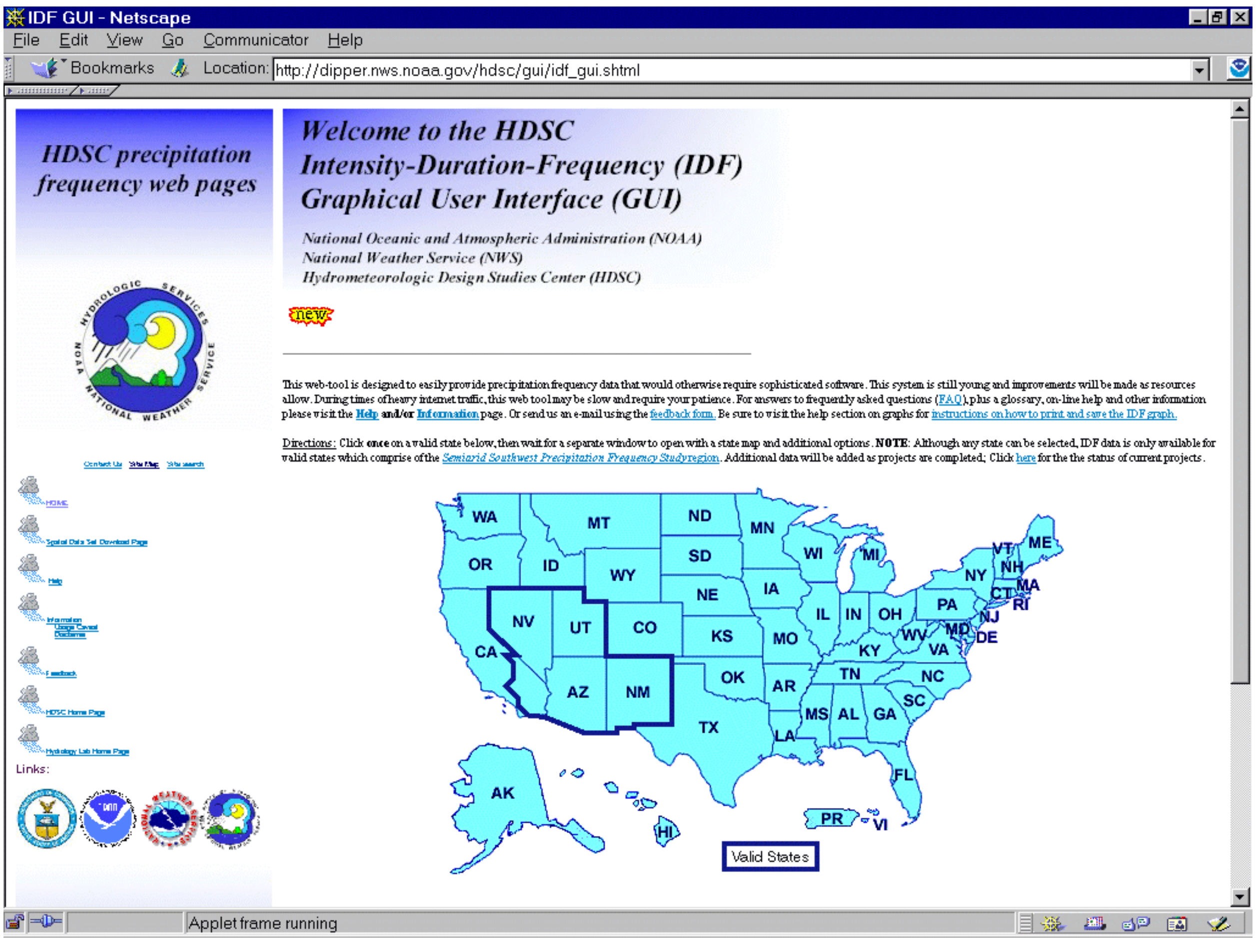


Figure 2. HDSC IDF GUI home page.

The GUI is based on a point-and-click interface, where the user chooses a point of interest from a shaded relief map complete with roads, cities and political boundaries. Figure 3 shows a portion of the shaded relief map in Arizona. Then duration (short [5min to 24 hours] or long [24 hours to 10 days]), units (inches or millimeters) and season (warm, cool or all) are selected. Based on these selections and the latitude/longitude pair, a color intensity-duration-frequency (IDF) curve (hyetograph) and data table are generated. Figure 4 shows the short duration IDF output for Flagstaff, Arizona. Both the table and graph are printable from a web browser, while the data is downloadable as text for further analysis.

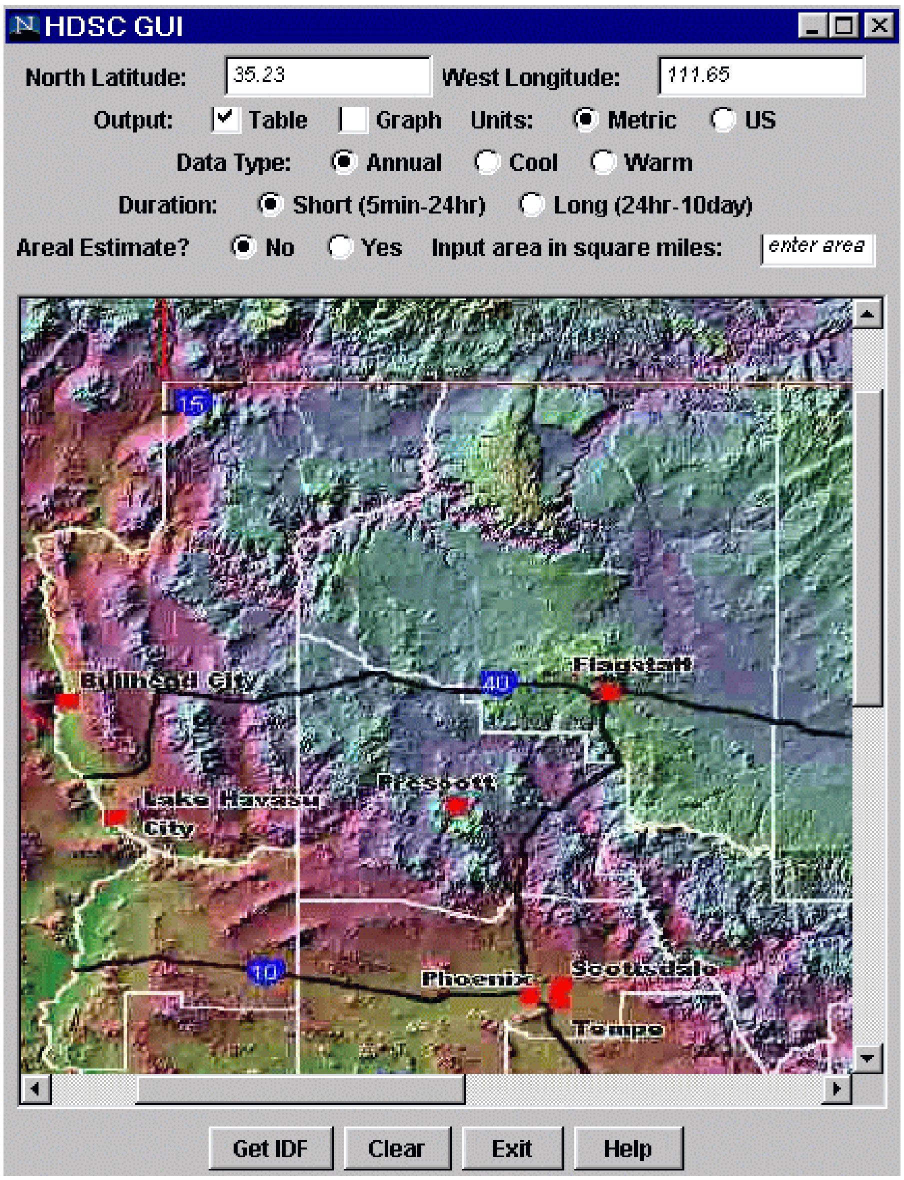
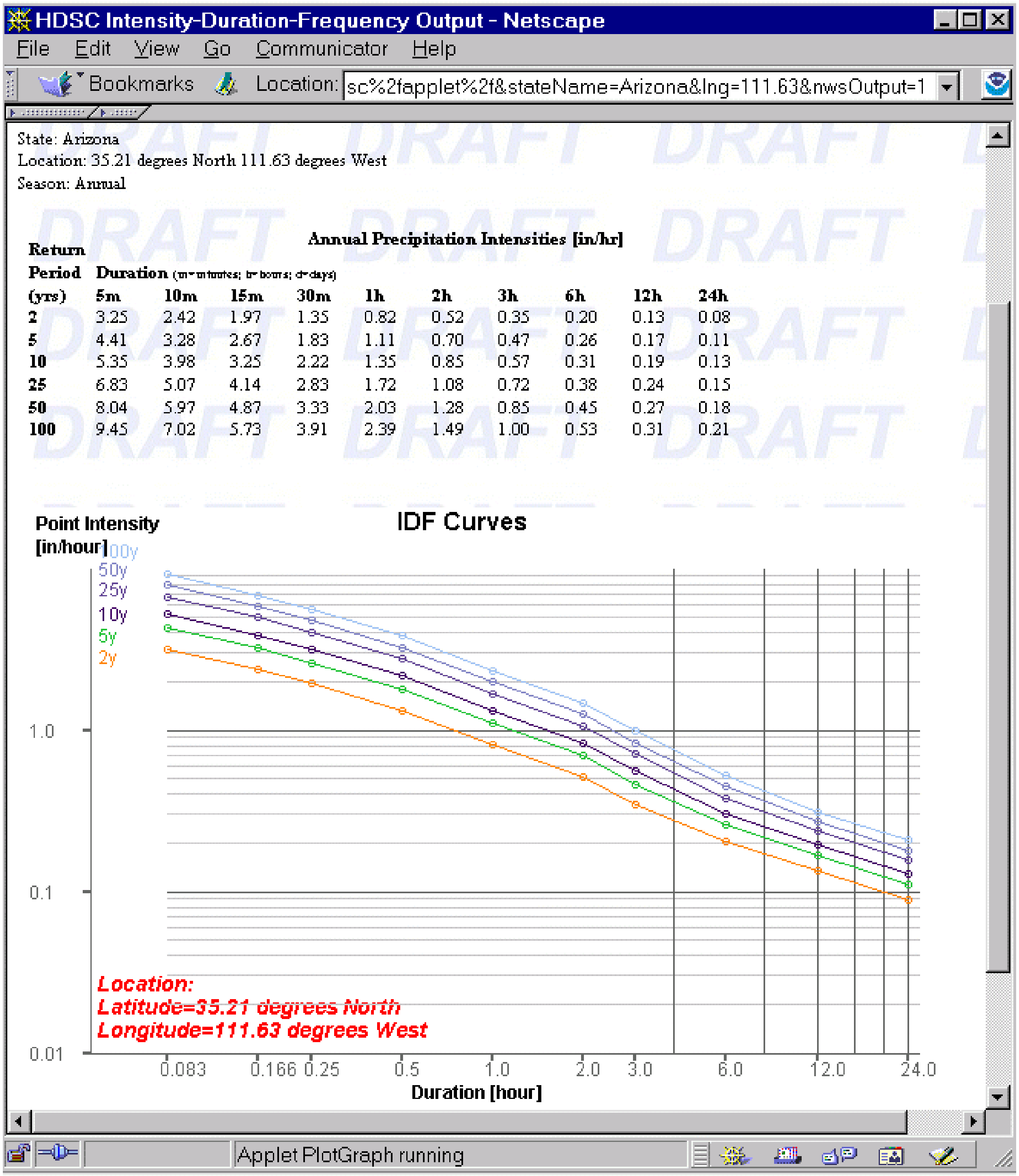


Figure 3. HDSC GUI shaded relief map and options window for Arizona.



State: Arizona
Location: 35.21 degrees North 111.63 degrees West
Season: Annual

Annual Precipitation Intensities [in/hr]

Return Period (yrs)	5m	10m	15m	30m	1h	2h	3h	6h	12h	24h
2	3.25	2.42	1.97	1.35	0.82	0.52	0.35	0.20	0.13	0.08
5	4.41	3.28	2.67	1.83	1.11	0.70	0.47	0.26	0.17	0.11
10	5.35	3.98	3.25	2.22	1.35	0.85	0.57	0.31	0.19	0.13
25	6.83	5.07	4.14	2.83	1.72	1.08	0.72	0.38	0.24	0.15
50	8.04	5.97	4.87	3.33	2.03	1.28	0.85	0.45	0.27	0.18
100	9.45	7.02	5.73	3.91	2.39	1.49	1.00	0.53	0.31	0.21

Duration (m=minutes; h=hours; d=days)

Figure 4. IDF output for short duration point estimates for Flagstaff, Arizona.

In addition to the IDF GUI, numerous supporting web pages have been developed to provide users with on-line help, information, downloading instructions and more. The help section includes definitions, frequently asked questions (FAQs), how areal estimates are calculated, derivation of maps, how a point estimate is derived, how to import spatial data into a Geographic Information System (GIS) and a search option. The information section includes details about the Java-based software, points of contact, usage caveat, references, sources and acknowledgments. Most importantly, the spatial dataset page allows users to download GIS-compatible raster maps. Links to Federal Geographic Data Committee (FGDC) compliant metadata and post-script files are included in the web pages, and eventually image (i.e., JPEG, GIF) files will be available. This will become the official Internet site for downloading precipitation frequency data for the United States.

The GUI and its associated web pages are installed and maintained on a new, dedicated web server at the National Weather Service headquarters in Silver Spring, Maryland.

References

Hershfield, D.M., 1961: Rainfall frequency atlas of the United States for durations from 30 minutes to 24 hours and return periods from 1 to 100 years, *Technical Paper No. 40*, U.S. Weather Bureau, Washington, DC.

Hosking, J.R.M. and J.R. Wallis, 1997: Regional Frequency Analysis, An approach based on L-moments, Cambridge University Press, NY, 224 pp.

Lin, B. and J.L. Vogel, 1993: A comparison of L-moments with method of moments, *Engineering Hydrology Symposium Proceedings*, ASCE, San Francisco, CA, July 25-30, 1993.

Wallis, J. R., 1993: Personal communication.

Appendix

The Appendix contains information about L-moments analyses.

METHOD OF L-MOMENTS

a) robust, less sensitive to sampling errors and outliers.
b) capable of characterizing a wide range of distributions.
c) linear combination of order statistics.

L-MOMENTS - DEFINITIONS

Random variable X with cdf F(X) and quantile func X(F):

$$X_{1:n} \le X_{2:n} \le \cdots \le X_{n:n}$$

L-MOMENTS for r = 1, 2, · · · are:

$$\lambda_r = r^{-1} \sum_{k=0}^{r-1} (-1)^k \binom{r-1}{k} E\, X_{r-k:r}$$

$$\text{L-CV} = \lambda_2/\lambda_1$$

$$\text{L-SKEWNESS} = \lambda_3/\lambda_2$$

$$\text{L-KURTOSIS} = \lambda_4/\lambda_2$$

L-MOMENT TESTS

1. DATA SCREENING--DISCORDANCY D_i

Let $\bar{U}_i = \left[t_2^{(i)}, t_3^{(i)}, t_4^{(i)}\right]^T$ be a vector for site i

$$\bar{U} = N^{-1} \sum_{i=1}^{N} \bar{U}_i \quad \text{-- unweighted group mean}$$

The Discordancy for site i is defined by:

$$D_i = \frac{1}{3}(\bar{U}_i - \bar{\bar{U}})^T \bar{S}^{-1} (\bar{U}_i - \bar{\bar{U}})$$

$$\text{where} \quad \bar{S} = (N - 1)^{-1} \sum_{i=1}^{N} (\bar{U}_i - \bar{\bar{U}})(\bar{U}_i - \bar{\bar{U}})^T$$

2. HETEROGENEITY--H

HOMOGENEOUS REGION - HETEROGENEITY, H: "IS THE BETWEEN-SITE DISPERSION OF THE SAMPLE L-MOMENTS FOR THE GROUP OF SITES LARGER THAN WOULD BE EXPECTED FROM A HOMOGENEOUS REGION"

[(OBS. DISP) - (MEAN DISP BY SIMULATION)]/SD OF SIM DISP

SIMULATION: 4-PARAMETER KAPPA DISTRIBUTION

$$H = (V - \mu_V)/\sigma_V$$

Where V Is the weighted SD at each site.

SAMPLE L-CV OR L-SKEW OR L-KURTOSIS

μ_V, σ_V FROM MONTE CARLO

3. GOODNESS-OF-FIT Z

$$Z^{GEV} = \frac{(\bar{t}_4 - \tau_4^{GEV})}{\sigma_4}$$

Where $\bar{t}_4$ is the regional average L-Kurtosis, τ_4 is the L-kurtosis of the fitted GEV distribution, and σ_4 is the standard deviation of τ_4 of the generated precipitation time series from Monte Carlo simulation.